

한갓 구현이 아닌 구성적 구조

The Compostional Structure without a Mere Implementation

최 훈[†]
Choi Hoon

요 약

포더와 그의 동조자들은 연결주의 모형이 구성적 구조를 보여주지 못하고, 또 보여줄 수 있더라도 자신들의 고전주의 모형을 물리적인 차원에서 구현하는 이론일 뿐이라고 공격한다. 연결주의가 구성적 구조를 보여주지 못한다고 주장할 때 중요하게 쓰이는 개념이 '고전적 요소성'이다. 그런데 글쓴이는 구성적 구조를 보이기 위해서는 고전적 요소성이 필요 없으며, 연결주의 모형이 고전적 요소성을 갖추고 있지 못하기 때문에 고전주의의 한갓 구현 노릇에서 빠져 나올 수 있다고 주장하겠다.

주제어 연결주의, 고전주의, 구성적 구조, 고전적 요소성, 구현

ABSTRACT

Fodor and his followers argue that connectionist model does not have compositional structure, or else it is merely an implementation of a classicist model. 'Classical constituency' is used importantly to show connectionist model does not have compositional structure. However, in contrast Fodor and his followers, I maintain that it is possible for connectionist model to have compositional structure without attributing classical constituency to cognitive model, and the very classical constituency makes the connectionist model nonimplementational.

Keyword Connectionist model, Classicist model, Compositional Structure, Classical Constituency, Implementation

[†]서울 대학교 철학과
Department of Philosophy,
Seoul National University

정신의 상태와 정신의 작용에 대한 고전주의 이론(규칙 지배적 기호 조작 이론)을 갈음하려는 새로운 이론을 표방하는 연결주의 모형은 포더와 필리신의 주목할 만한 비판(Fodor/Pylyshyn, 1988)의 물결에 맞서 오히려 그 위상이 한껏 높아진 느낌이다. 의미있는 인지 이론이라면 우리 정신의 상태와 작용의 부정할 수 없는 현상인 구성적 구조를 설명할 수 있어야 하는데 연결주의 모형은 그렇지 못하기에 인지 이론의 지위에 오르지 못한다고 하는 포더와 필리신의 여문 지적 이후, 연결주의를 인지 이론으로서 옹호하는, 나아가 거기에 열광해 하는 이들은 그 지위를 얻기 위해 연결주의 모형도 구성적 구조를 보여줄 수 있음을 증명하는 연구들을 내놓고 있다((1), (2), (3), (4), (5), (6), (7), (8)). 글쓴이도 (9)를 통해 그 대열에 끼어 들었다. 그 글에서 의식하지 못한 것은 아니지만 그리고 앞으로의 연구 과제로 남겨 놓긴 했지만, 우리는 포더와 필리신이 요구하는 구성적 구조를 갖추었다고 손을 털고 안심하고 일어났을 때 그들이 내놓은 또 하나의 공격이 기다리고 있다는 것을 잊지 말아야 한다. 바로, 그런 구성적 구조를 보여줄 수 있는 모형이 되더라도 연결주의 모형은 인지 차원의 고전주의 모형을 물리적으로 구현하는 이론일 뿐이라는 공격말이다. 고전주의를 대치하려는 연결주의자들의 노력이 고전주의자들을 도와주는 결과만 될 뿐이고 새로운 인지 이론이라는 시도가 애초에 헛된 것인지도 모른다.

글쓴이는 위 글에서 포더와 필리신의 첫번

째 공격을 넘어설 수 있는 길로 스몰렌스키의 텐서 곱 표상을 보기로 들었다. 그러나 포더와 맥래플린은 과연 이 모형이 포더와 필리신의 공격을 피해갈 수 있는 답변이 될 수 있는지 심각하게 의심하고 있다((10)). 그들의 의심에서 중요한 구실을 맡고 있는 개념이 '고전적 요소성'이다. 즉 텐서 곱 표상의 요소는 실제로 사례화하지 않는 요소이기에 인과적인 유효성을 가지지 못하고 따라서 구성적 구조를 이루는 참다운 요소가 아니라는 것이다. 그런데 나는 오히려 '고전적 요소성' 개념이 연결주의가 구현의 덩을 벗어날 수 있는 길이 아닐까 생각해 본다.

결론부터 말하자면 이렇다. 포더와 맥래플린은 자신들이 구성적 구조 논의에서 꼭 필요하다고 생각하는 고전적 요소성을 스몰렌스키의 텐서 곱 모형은 지니고 있지 못하기에 그 모형은 포더와 필리신에 대한 답변으로서 실패한다고 말하고 있지만 나는 구성적 구조를 보이기 위해서는 고전적 요소성이 필요 없으며 연결주의 모형에서 말하는 요소도 인과적으로 유효하다는 것을 논의하고, 연결주의 모형이 고전적 요소성을 갖추고 있지 못하기 때문에 오히려 고전주의의 한갓 구현 노릇에서 빠져 나올 수 있다고 주장하겠다. 그러기 위해서 구현의 조건을 되도록 엄격히 하겠다.

먼저 다음 두 절에 걸쳐 포더와 필리신의 연결주의에 대한 두 가지 공격을 들어보고, 연결주의 모형은 구성적 구조를 왜 못 보이는지 알아보자.

1. 포더와 필리신의 연결주의 공격

포더와 필리신은 [11]에서 새로운 인지 모형 이고자 하는 연결주의자를 다음과 같은 딜레마에 빠뜨린다.

- (1) 인지 체계는 X를 가져야 한다. 따라서
- (2) 연결주의 체계는 (1) X를 가지지 못하거나—이 경우에 연결주의는 적합한 인지 모형이 되기에는 너무 약하고, 아니면 (2) 연결주의 체계는 단지 고전적인 체계(기호 조작 모형)의 구현에 불과하다—이 경우에 연결주의는 인지에 대한 새로운 개념이 되지 못한다.

연결주의가 인지 체계에 대해 무언가 새로운 얘기거리를 내세운다고 생각한다면 그 이론은 자신들이 인지 체계가 꼭 가져야 한다고 가정하는 어떤 본질적인 속성 X를 가지지 못하는 한갓 연상주의(associationism)에 불과하다. 또 연결주의가 그런 X까지 포섭할 수 있을 정도로 의미있는 이론이고자 한다면 그 이론은 더 이상 새로운 이론이 아니며 자신들이 옹호하는 고전주의 이론의 한갓 구현(implementation)에 불과하다.

이 때 포더와 필리신이 X로서 들고 있는 것은 바로 표상의 구성적 구조이다¹⁾. 한 체계가 구성적 구조를 가졌다는 것은 그 체계에 기초가 되는 요소가 있고 그 요소들이 요소 표현들을

이루며 복합 표현은 그 요소 표현들로부터 되부름적(recursive) 규칙에 의해 결정된다는 것을 말한다. 곧 복잡한 표현의 내용은 그 복잡한 표현을 구성하고 있는 (상대적으로 더) 간단한 부분의 내용에 의해 결정된다는 것이다. 포더와 필리신은 이런 구성적인 구조를 표상은 꼭 가지고 있어야 한다고 생각하는데 왜냐하면 그래야만 우리 표상에서 부인할 수 없는 현상인 생산성과 체계성을 설명할 수 있기 때문이다. 그런데 표상을 다루는 인지 모형이고자 하는 연결주의에서는 표상이 이 구성적 구조를 지니고 있지 못하다. 그러므로 연결주의는 적절한 인지 모형이 되기에는 약하다는 것이 포더와 필리신이 내세우는 바이다. 그런데 이런 구성적인 구조를 연결주의가 보여준다고 하더라도 그것이 고전적인 인지 구조를 구현하는 방법

주1) 차머스는 표상의 구성적 구조가 표상이 가지는 본질적 속성 X의 한 측면일 뿐임을 강조한다. 그래서 포더와 필리신이 구성성이 X의 전부 다인 것처럼 생각하는 것은 잘못된 전체라고 말한다. 마음의 한 측면일 뿐인 그 구성적 구조를 연결주의는 못 보이고, 고전주의는 보였다고 해서 고전주의가 마음에 대한 최선의 이론이라는 결론을 내릴 수도 없으며, 연결주의가 구성성의 고전적 이론을 구현할 수 있다는 사실은 연결주의가 마음에 관한 고전적 이론을 구현한다는 것을 함축하지도 않는다고 한다([2]). 차머스의 주장은 옳은 점도 있다. 그러나 그 주장이 포더와 필리신의 주장을 제대로 이해했느냐도 문제겠지만([12]), 구성적 구조가 마음이 가지는 한 측면일 뿐이라고 해도 중요한 측면임에 모두 동의한다면 그것을 설명해야 하는 의무를 연결주의가 벗어버리지는 못할 것이다.

을 통해서라면 말짱 헛일이다. 그렇다면 그 이론은 인지 차원에서 주장되는 고전주의 인지 이론에 대한 물리적인 구현 차원의 노릇들이 될 뿐이므로 전혀 인지에 대한 고전주의 이론을 뒤집을 수 있는 새로운 이론이 되지 못하기 때문이다. 요컨대 연결주의 이론이 고전주의 이론의 대안이 되면서 적절한 이론이라면 고전주의 인지 구조의 한갓 구현이 되지 않으면서 구성적 구조를 설명할 수 있어야 한다는 것이 포더와 필리신의 주장인 것이다.

맥래플린을 따라 위 딜레마의 (a)와 (b)를 각각 **적합성 조건**과 **비구현 조건**이라고 부르고 ((13)) 각 조건을 더 들여다 보자.

(ㄱ) 적합성 조건²⁾

우리 자연 언어가, 더 엄격히 말하자면 자연 언어의 구문론과 의미론이 구성적이라는 것은 일반적으로 받아들여지고 있다. 다시 말해서 자연 언어의 일반적인 구문론과 의미론은 기초가 되는 요소들과, 그 요소들로부터 이루어지는 더 복잡한 표현을 결정하는 (아마도 되부름적인) 규칙들로 이루어져 있는 것이다. 이를테면 '성자는 귀남이를 사랑한다'는 표현은 구성적인데 그 표현은 '성자', '귀남', '사랑한다'를 그 요소로 지니고 있고 그 요소들에 의해 '성자는 귀남이를 사랑한다'고 하는 복합 표

현을 만든다. 이렇게 볼 때 구성성에서 **요소** (constituency)가 중요한 개념임을 알 수 있다. 복합 표현이 요소 표현으로 환원되고 이 요소 표현으로부터 복합 표현을 만들 수 있으며 그 과정이 되부름적인 규칙에 의해 일반적이고 믿을 만하고 효율적인 방법으로 이루어진다는 것이 구성성 개념의 중요한 축이다.

자연 언어가 구성적인 구문론과 의미론을 갖는다는 사실은 보통 **생산성 논변**(productivity argument)을 통해서 설명된다. 그 논변은 다음과 같다. 화자는 무한한 수의—아니면 적어도 한정없이 큰 수의—새로운 문장들을 이해할 수 있다. 그렇게 무한한 수의 문장을 이해한다는 것을 어떻게 설명해야 할까? 이에 대해서는 두 가지 대답이 가능하다. 첫째는 무한한 수의 문장 각각에 해당하는, 곧 그 문장이 뜻하는 무한한 수의 사실들이 있다고 가정하는 것이다. 이 무한한 사실들을 묶어 설명해 내는 유한한 토대가 되는 사실들은 있지 않다. 있는 것은 오직 무한한 사실들뿐이다. 그러나 이 설명은 그럴 듯하지 못하다. 자연 언어를 쓰는 우리가 항상 새로운 문장에 부딪칠 때 그 문장을 이해해 내고 설명해 내며 또 새로운 문장을 생산해 내는 유한한 토대가 존재하지 않는다면 우리가 그 문장을 이해하고 학습한다는 것은 신비일 뿐이다. 그렇다면 우리는 다른 대답을 기다릴 수밖에 없다. 곧 우리가 무한히 많은 문장들을 이해할 수 있는 것은 우리가 문장을 이해한다는 것이 결국은 그 문장들의 부분들, 요소들을 이해하는 것이기 때문에 가능하다는 설명이다.

주2)(9), 15-20쪽 참조.

한 문장이 몇 개의 요소들로 이루어졌다고 하면 서로 다른 문장들은 구성 요소들의 종류와 이 요소들을 정렬하는 서로 다른 방식에 상응하게 된다. 그러면 새로운 문장은 새로운 정렬 방식과 상응하게 되는 것이다. 그러므로 한 문장의 의미는 그 요소 구조에 의해 결정된다. 이는 무한한 문장들이 주어졌을 때 좀더 유한한 요소들을 통해서 이해할 수 있고 또 이 요소들을 통해서 무한한 새로운 문장들을 만들어 낼 수 있다는 생각이다. 이렇게 봤을 때 생산성은 구성성에 의존하고 이 구성성에서는 요소성이 중요한 개념임을 알 수 있다. 포더의 지적처럼 “한 대상이 요소를 갖지 않는다면 그 대상의 의미를 요소의 의미로부터 구성해낼 수 없다” ([14], 150쪽).

이상의 논변은 다음과 같은 구조를 갖는다.

- (1) 우리 자연 언어에는 무한히 많은 문장들이 있고 우리는 그 문장들을 이해한다.
- (2) 우리 자연 언어가 구성적인 구조를 가지면 (1)이 잘 설명된다.
- (3) 그러므로 우리 자연 언어는 구성적인 구조를 갖는다.

(1)은 우리 언어의 생산성이라는 경험적 사실이다. 구성성은 그 경험적 사실을 설명하기 위해 요청되는 원리이다. 우리 언어에 구성적 구조를 가정하면 언어의 생산성이라는 경험적 사실을 잘 설명할 수 있다는 게 이 논변의 구조이다.

포더와 필리핀은 자연 언어의 구성적 구조를 받아들이게 하는 근거가 되는 (1)의 생산성이라는 현상이 너무 강한 주장이어서 그걸 받아들이지 않는 이들이 있을까 조심스러워한다. 그래서 우리 자연 언어에서 발견할 수 있는 다른 현상, 그러면서 어느 누구나 동의하리라고 생각하는 현상, 곧 체계성(systematicity)도 구성적 구조를 지지할 수 있는 근거라고 내세운다. 언어 능력이 ‘체계적’이라는 말로 뜻하는 바는 어떤 문장을 말하거나 이해하는 능력은 다른 많은 관련된 문장들을 말하거나 이해하는 능력과 본질적으로 연관되어 있다는 것이다 ([11], 37쪽; [14], 149쪽). 이를테면 “성자는 귀남이를 사랑한다”를 말하고 이해할 줄 아는 이가 “귀남이는 성자를 사랑한다”를 말하고 이해할 줄 모를 리가 없다. 이 체계성도 우리 언어가 구성적인 구조를 가졌기에 설명할 수 있다. “성자는 귀남이를 사랑한다”는 문장을 ‘성자’, ‘귀남’, ‘사랑한다’의 요소로 분해할 수 있고 그 요소를 통해 다시 “귀남이는 성자를 사랑한다”를 만들 수 있기에 “귀남이는 성자를 사랑한다”는 문장을 이해할 수 있는 것이다. 결국 체계성도 구성성을 요구하게 되고 구성성에서는 요소라는 개념이 중요한 구실을 함을 다시 한 번 확인할 수 있다. 그런데 고전주의 인지 이론에서 이 요소에 어떤 조건을 덧붙이지는지는 그리고 그것이 고전주의의 구성성 개념에 얼마나 중요한지는 조금 있다. 포더와 맥래플린의 주장을 들으면 깨달을 수 있을 것이다.

이상과 같은 논변은 그 수많은 자연 언어 문

장들이 표현하는 사고에도 그대로 적용할 수 있다. 즉 사고도 구성적인 의미론을 틀림없이 갖는다((14), 148쪽).

(L) 비구현 조건³⁾

그러나 여전히 딜레마의 두번째 뿔이 남아 있다. 연결주의가 설사 구성적인 구조를 보이는 모형이 된다고 하더라도 연결주의는 참된 인지 과정에 대한 설명이 아니라 구현 차원에 대한 설명일 뿐이므로 인지 이론과 상관없는 차원의 이론일 뿐이라는 것이다. 포더와 필리신은 인지 현상에 대한 의미있는 이론은 일정한 기술의 차원에서 성립하는데 그 차원은 풀이법(알고리즘)과 표상의 차원이라고 한다. 그런데 연결주의가 고전주의보다 우월하다고 알려진 점들((9), 140-1쪽)은 모두 구현 그리고 이론가들이 특정 경우에 쓰곤 하는 풀이법의 특정 실현(realization)과 관련된 특성들일 뿐이고 그러기에 인지 이론과는 상관없는 차원에서 말해지는 이론일 뿐이라고 주장한다((11), 65쪽).

포더와 필리신의 말이 옳다면 연결주의 모형은 결코 인지과학에 대해 혁명적인 새로운 토대를 제공한 것이 되지 못한다. 그것은 기껏해야 인지체계의 구현 이론일 뿐이다. 다시 말하면 일정한 정보 처리 체계의 추상적 구조가 아니라 그 물질적 기반을 다루는 이론일 뿐이다. 인지론자들에 따르면 우리가 심리학이라 할 만

한 것은 일정한 수준의 기술의 차원(level of description)에서 성립할 수 있는 것인데 그러한 기술의 차원에서만 우리가 관심을 가질 수 있는 행위에 대한 설명과 예측이 언어될 수 있다. 그런데 포더와 필리신은 그런 기술의 차원은 구현의 차원이 아니라 풀이법과 표상의 차원이고 연결주의는 구현의 차원에 있는 이론이라고 보는 것이다. 더구나 포더와 필리신은 단순히 풀이법과 표상의 차원에서만 의미있는 심리학 이론이 가능하다고 주장하는 데 그치는 것이 아니라 그 풀이법과 표상은 앞서 X에 해당하는 특성들을 갖는 것이어야 한다고 생각하는 듯하다. 따라서 연결주의는 기본적으로 구현의 차원에서 주장되는 것이므로 심리학적인 의미를 갖추고 있는 모형이라고 할 수 없을 뿐더러 설사 그 구현의 차원이 풀이법/표상의 차원을 수반해서 갖는다고 하더라도 연결주의는 표상에 대한 비기호적인 모형을 갖고 있기에 포더와 필리신의 입맛에 맞지 않는 모형인 것이다.

여기서 우리가 열른 알 수 있는 것은 포더와 필리신이 이론의 대상이 되는 자연을 조직화할 때 그 조직화가 몇가지 차원(level)으로 구분된다는 전제를 갖고 있다는 점이다. 사실 문제가 되는 것은 누구나 할 수 있는 그런 전제 자체는 아니다. 그 차원 구분이 어떤 근거로 이루어지고 있고 또 차원들 사이의 구분에 어떤 태도를 보이고 있느냐 하는 점이다. 먼저 포더와 필리신은 차원에 의한 설명을 마아(Marr)가 인지과학에 대해 했던 영향력있는 분석에 토대를 두

주3) (15) 참조.

고 있는 듯 하다. 마이는 인지과학의 설명을 세 가지 차원으로 나눈다.

- a **연산 차원:** 연산의 목적은 무엇인가? 그것이 왜 적합한가?
- b **표상과 풀이법 차원:** 이 연산이론은 어떻게 구현될 수 있는가? 특히, 입력, 출력에 대한 표상은 무엇인가? 그리고 풀이법에 대한 풀이법은 무엇인가?
- c **하드웨어 구현 차원:** 표상과 풀이법은 물리적으로 어떻게 실현되는가?

앞에서도 설명했지만 되풀이하자면 이 가운데 심리학 연구가 문제되는 차원은 표상/풀이법 차원이다. 효율성, 소음이나 물리적 손상에서 퇴화의 문제, 특정 문제가 쉬운가, 어려운가, 어느 문제가 빨리 풀리고 푸는 데 얼마나 시간이 걸리는가, 정보가 어떻게 표상되는가 하는 문제들에 관심을 갖는 것은 바로 표상/풀이법 차원 예시이다([16], 122-3쪽).

자연을 조직화할 때 그 조직화가 몇 가지 단계로 구분된다는 전제는 그리 새로운 것도 없는, 그래서 큰 시비거리가 될 것 같지는 않다. 그 전제는 자연이란 서로 다른 크기의 존재물(entity)들로 이루어져 있고 그 중 더 작은 존재물은 더 큰 존재물의 구성 성분(component)이라는 생각이다. 우리는 물론 각 차원 안에서 존재자들의 상호작용에 대해 인과적인 이야기를 할 수 있다. 이를테면 분자들 사이의 상호작용에 대한 이야기, 강 또는 산 사이의 상호작용

에 대한 이야기, 우주 사이의 상호작용에 대한 이야기... 식으로 말이다. 그런데 문제는 그 다음에 생긴다. 서로 다른 차원에 있는 존재자들끼리의 상호작용에 대해서는 어떤 태도를 취하고 있는나 하는 점에서 포더와 필리신은 그들의 자세를 드러낸다. 그들은 각 차원에서 할 수 있는 이야기들이 서로 독립적이라고 주장한다.

과학자들이 이런 차원들 가운데 하나에서 세계가 갖는 인과 구조에 대해 하는 이야기는 그 위 또는 아래 차원에서 인과 구조에 대해 하는 이야기와 완전히 다르다([11], 9쪽).

서로 다른 차원들에서 하는 과학적 설명은 서로를 최소한으로만 제한한다. 따라서 차원들 사이에서 얘기되는 인과적 이야기란 없다. 포더와 필리신이 보기에 인지 이론에 적합한 인과적 이야기란 표상, 그것도 기호적 표상에서 수행되는 행동들에 대한 이야기이다. 연결주의만으로는 표상에서 이루어지는 행동들에 적절한 이야기를 해줄 수 없고, 기호적 설명만이 그것을 할 수 있으므로 연결주의는 인지 이론의 후보로 고개도 내밀지 못하는 것이다.

이렇게 본다면 연결주의 모형은 인지 과정의 본성에 대해 원리적으로 중립적이다. 그 모형은 기껏해야 뇌(또는 아마도 이상화된 뇌와 같은 연결망)가 전통적인 인지과학이 가설화한 과정의 유형들을 현실화한 방법을 설명하려고

시도하여, 고전적인 정보 처리 심리학의 목표를 진보시켰다는 평가를 받을 수 있을 뿐이다. 맥 래플린은 한 술 더 떠 그러면 연결주의에도 더 좋은 일이 아니겠냐고 말한다. “도대체 구현에 ‘한것’이라는 낱말이 왜 붙는가?” ([13], 184 쪽).

유기체의 비표상적 상태들이 연결주의 네트워크를 구성한다고 말하는 것은 아무런 쓸모가 없다. 왜냐하면 그렇게 말하는 것은 마음이 심리적 차원, 곧 표상차원에서 어떤 연결망이냐고 묻는 질문에 아무 것도 대답하는 것이 없기 때문이다. 연결주의 모형이 심리적 차원의 설명에 아무런 도움을 줄 수 없는 것은 분자 모형이나 양자 역학 모형, 생화학 모형이 그럴 수 없다는 것과 마찬가지로이다.

2. 구성적 구조가 없는 연결주의 모형

자연 언어의, 그리고 표상의 구성적 구조를 앞 절에서와 같이 이해했다고 할 때 포더와 필리신은 연결주의 모형이 그런 구성적 구조를 보이지 못한다고 주장한다.

연결주의는 유닛들 사이의 기본적인 관계로서 인과적인 관련성만을 인정한다. 만약 당신이 유닛들 사이에 흥분과 억제 어떻게 일어나는가 알고자 할 때 네트워크 안에 있는 유닛들이 서로 어떻게 관련을 맺는가에 대해서만 알면 다 된다. 이에 비해 고전적인 이론은 의미론적인 가치를 부여

한 대상들의 인과적인 관계뿐만 아니라 구성성이 성립되는 구조적인 관계까지도 인정한다([11], 12쪽).

결론은 간단하다. 자신들이 주장해온 고전주의 모형은 구성적인 구조를 보이고 있지만 연결주의는 구성적 구조를 보이지 못한다는 것이다. 그게 두 이론 사이의 차이점이고 그 차이점은 연결주의에서 의미있는 인지 이론의 지위를 뺏을 수 있을 만큼 큰 것이다. 연결주의는 왜 구성적 구조를 보이지 못한다고 생각하는가? 연결주의 체계에서는 개별적인 유닛이나 유닛의 연합이 의미론적으로 해석될 수 있다. 그러나 유닛들 사이의 기본적인 관계는 인과적인 관계만을 인정한다. 이 인과적인 관계에 대해서 포더와 필리신은 자세한 설명을 하고 있지 않으나 짐작컨대 유닛과 유닛 사이에 연상적인 관계만을 인정하고 부분과 전체의 관계일 때 나올 수 있는 포함과 도출 관계는 전혀 고려하지 않는 관계이다. 있는 것이라고는 오직 유닛 사이에 미치는 힘—고전적인 연상주의에서는 빈도, 연결주의에서는 활성화 정도—뿐이다. 모든 유닛은 각각이 지니는 그리고 서로에 미치는 힘을 제외하고는 구분이 되지 않는 원자적인 상태이다. 또 구문론적 규칙을 적용하려면 적당한 요소가 있어서 거기에 되부름적인 규칙을 쓸텐데 유닛 하나하나가 모두 원자적인 상태이니 구문론적 규칙에 따라 조작할 나위도 없다([9], 30-3쪽 참조).

이와 같은 포더와 필리신의 비판에 대해 연

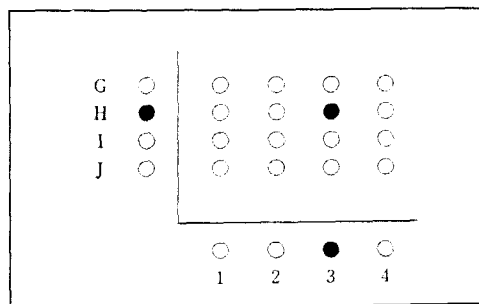
결주의자의 처지에서는 개념적인 답변과 경험적인 답변이 가능하다. 여기서는 경험적인 답변, 그 중에서도 특히 스몰렌스키의 텐서 곱 표상을 포더와 필리신 논변에 대한 반례로서 내놓으려고 한다.⁴⁾ 경험적인 답변이 철학자들에게는 어떻게 보면 더 강력해 보이고, 또 포더와 맥래플린이 이 반례에 대한 해석을 놓고 맞대응을 하는 것이 이 글의 4절과 5절의 논의거리가 되기 때문이다.

**3. 연결주의자의 답변:
스몰렌스키의 텐서 곱 표상**

스몰렌스키의 근본적인 생각은 구성적 구조를 벡터 연산으로 표상하려는 것이다. 그 과정을 (4)에 제시된 생각을 따라 가면서 "성자는 귀남이를 사랑한다"고 하는 구성적 구조를 어떻게 연결주의 네트워크가 표현해 내는지 알아보자.

먼저 익혀야 할 개념들이 있다. 스몰렌스키의 텐서 곱 표상에서는 구조화된, 곧 구성적인 구조를 갖는 대상의 표상은 '특정 역할(role)에 묶인(bound) 점유자(filler)', 다시 말해 특정 역할을 맡은 점유자로 보고 있다. 이것을 구체적으로 알아보기 위해 맥클러랜드와 루멜하트의 낱말 지각 모형((17))에 적용해 보자.

'JOHN' 과 같은 임의의 녀 자짜리 낱말을 표상한다고 생각해 보자. 이 때 임의의 낱말에서 각 글자가 위치를 점유할텐데 그 때 글자를 표상하는 데 쓰이는 유닛(unit)들을 **점유자 유닛**이라 부른다. 'JOHN' 과 같은 경우는 그 단어가 'J', 'O', 'H', 'N' 에 의해 점유되어 있으므로 'J', 'O', 'H', 'N' 각각이 점유자 유닛이 될 것이다. 그리고 글자의 위치들은 글자가 행하는 역할이라고 생각할 수 있으므로 그 위치들을 표상하는 **역할 유닛**을 생각해 볼 수 있겠는데 이 경우는 각 위치-첫번째, 두번째...-가 역할 유닛들이 될 것이다. 그리고 마지막으로 이 두 가지 유닛들을 통해서 우리는 특정 글자를 표상할 수 있다. 곧 첫번째 위치에 있는 'J' 하는 식으로 말이다. 그것을 표상하는 데 쓰이는 유닛이 **묶음 유닛(binding unit)**이다. 곧 이 유닛은 **역할에 묶인 점유자(a filler bound to a role)**를 표상한다.

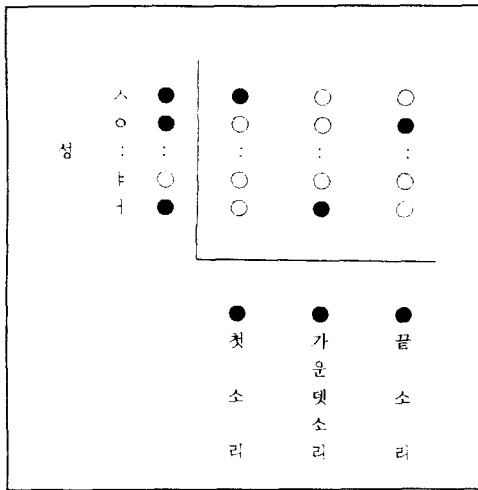


(그림 1) 세번째 자리에 있는 'H'의 표상

이해를 돕기 위해 '세번째 자리에 있는 H'를 표상하는 것을 (그림 1)로 나타내 보자. 이

주4) 개념적인 답변에 대해서는 (9), 2장 참조.

그림에서 왼쪽 가장자리에 있는 유닛이 점유자 유닛이고 맨 아래쪽의 유닛이 역할 유닛이며 나머지 유닛들은 묶음 유닛이다. 여기서 검은 동그라미는 유닛이 활성화되었음을 뜻한다.



(그림 2) 글자 '성'의 표상

이제 텐서 곱 개념을 끌어들이 때이다. 점유자 f 를 역할 r 에 묶는다고 하자. 이 때 역할 r 의 표상은 벡터 공간 V_R 에서의 벡터 r 이다. 또 점유자 f 의 표상은 벡터 공간 V_F 에서의 벡터 f 이다. 그러면 f/r 이 역할 r 에 점유자 f 를 묶음을 뜻한다면 f/r 의 표상은 텐서 곱 벡터 공간 $V_B = V_F \otimes V_R$ 에서의 **텐서 곱** 벡터 $f/r = f \otimes r$ 이다. 위 글자 표상 모형에서 생각해 본다면 (그림 1)의 왼쪽 가장자리에 있는 유닛들이 한 활성화 벡터를 이루고 아래쪽에 있는 유닛들이 다른 활성화 벡터를 이룬다고 볼 수 있는데 일단 그렇게 본다면 이 벡터들의 텐서 곱 벡터-나머지 유닛

들-가 알맞는 위치에 있는 알맞는 글자를 표상한다고 주장할 수 있다.

이제 이 개념들을 이용하면 "성자는 귀남이를 사랑한다"의 구성적 구조를 표상하기 위해 먼저 '성', '자'.... 각 글자를 표상할 수 있다. 첫 글자 '성' 만 보기로 들어보면 (그림 2)와 같이 될 것이다.

이런 식으로 각 글자를 표상한 다음 그 글자를 통해 낱말을 표상하기 위해 우리의 시각을 더 넓혀야 한다. 곧 우리 네트워크에서는 주어진 위치에 있는 글자의 표상 뿐만 아니라 낱말 자체의 분산 표상까지 원한다. 이 때 도입되는 개념이 벡터 합, **패턴 중첩**(superposition)이다. 만약 두 글자가 연결주의 네트워크에서 어떤 활성화 패턴에 의해 각각 표상되었다면 이 두 글자를 접속(conjunction)한 복합 글자, 또는 낱말의 표상은 두 개별 패턴을 중첩시켜서 나온 패턴이다. 이것은 각 위치에 있는 각 글자에 해당하는 벡터들의 텐서 곱을 그냥 더하기만 하면 된다. 곧 f_{se} 과 f_{ja} 가 각각 글자 '성' 과 '자' 의 점유자 벡터이고 r_1 과 r_2 가 각각 첫번째 위치, 두번째 위치를 가리키는 역할 벡터라면 '성자'의 벡터값은

$$f_{se} \otimes r_1 + f_{ja} \otimes r_2$$

이다.

이렇게 하여 우리는 '성자', '귀남', '사랑한다'의 표상을 갖게 되었다. 이제 "성자는 귀남이를 사랑한다"의 표상을 이 요소 표상들로부터

터 만들어 낼 수 있다. '성자'와 '귀남'은 이 새로운 표상을 만들기 위해 새로운 점유자 유닛이 된다. 이런 점에서 벡터 곱, 또는 벡터 합은 되부름적 과정이다. 그러면 그 점유자 유닛과 결합하는 역할 유닛은 각각 '행위자'와 '피동자'가 될 것이다. 그래서 '행위자로서의 성자'의 표상은 $[성자 \otimes r_{행위자}]$ 라는 텐서 곱이 될 것이고 '피동자로서의 귀남'의 표상은 $[귀남 \otimes r_{피동자}]$ 라는 텐서 곱이 될 것이다. '행위자로서의 성자'와 '피동자로서의 귀남'을 복합시킨 표상을 만드는 것은 이제 어렵지 않다. 이 두 텐서 곱을 중첩, 벡터 합하면 될 것이다. 우리가 알고자 하는 최종적인 표상, 곧 "성자는 귀남이를 사랑한다"는 이 벡터 합에 '사랑한다'에 해당하는 벡터 $[값-아마도 [사랑한다] \otimes r_{행위}-을 더하면]$ 될 것이다. 곧 '성자는 귀남을 사랑한다'의 벡터 표상은

$$[성자 \otimes r_{행위자}] + [귀남 \otimes r_{피동자}] + [사랑하기 \otimes r_{행위}]$$

와 같이 될 것이다(9), 85-9쪽 참조).

이상과 같은 텐서 곱 표상은 포터와 필리신의 반론을 벗어날 수 있을 것 같다. 한 구조화된 대상의 텐서 곱 표상은 그 대상의 요소들 각각에 대한 점유자/역할 묶음을 표상하는 벡터들을 합해서 만들 수 있기 때문이다.

텐서 곱 표상에 대해서 두 가지 특징을 지적해야겠다. 첫째로 텐서 곱 표상은 단순한 부분/전체 관계는 아니라는 것이다. 초기의 국부적인

(local) 연결주의 모형에서처럼 특정한 유닛들이 특정한 고정된 표상 내용에 부여되기만 하고 그 외에 구조를 반영할 만한 아무런 장치를 갖추고 있지 못한 경우를 생각해 보자. 그렇다면 '성자', '귀남', '사랑한다'에 각각 해당하는 특정 마디(node)들이 있을 텐데 그 마디들이 동시에 발화한다고 할 때 '성자는 귀남이를 사랑한다'는 표상과 '귀남이는 성자를 사랑한다'는 표상을 구분할 방법이 도대체 없고 그래서 체계성이 설명 안 되는 것이다. 그 때의 전체 표상과 요소 표상은 단순히 부분/전체의 포함 관계일 뿐이고, 임자말[주어], 풀이말[술어] 등과 같은 풍부한 구문론적 요소 관계를 반영하지 못하는 것이다.⁵⁾ 그런데 이미 보았던 것처럼 텐서 곱 표상의 점유자/역할 묶음은 한 표상의 요소 표상을 나타내줄 뿐만 아니라 임자말, 부림말[목적어], 움직임[동사]와 같은 구조적 관계까지도 보여줄 수 있다.

또 다른 특징은 텐서 곱 표상을 풀어 헤쳐 그 요소 표상들을 되찾을 수 있다는 것이다. 곧 요소 표상을 통해 구조화된 복합 표상의 텐서 곱 표상을 합성할 수 있을 뿐만 아니라, "성자는 귀남이를 사랑한다"를 표상하는 벡터에서 '성자'와 '귀남', 그리고 '사랑하기'를 표상하는 벡터를 분명히 일반적이고 효율적이고 믿을 만한 방식을 통해서 찾을 수 있는 것이다.⁶⁾ 나

주5) [18], 206-7쪽 참조.

는 [9]에서 구성적 구조는 '그 요소를 복합 표현에서 분해해 내고 다시 복합 표현으로 합성해 낼 수 있는 일반적이고 효율적이고 믿을 만한 방식'을 만족하는 조건이어야 한다고 말했다 (19쪽). 텐서 곱 표상은 이 조건을 만족함을 알 수 있다.

4. 포더와 맥래플린의 의심

나는 앞절에서 텐서 곱 표상을 통해 연결주의의 표상도 요소 구조를 가질 수 있고 그러기에 구성적 구조를 보일 수 있다고 주장했다. 그런데 포더와 그의 동조자들은 연결주의에서 말하는 요소는 참다운 요소가 아니라고 주장한다.

주6) 스톨렌스키는 이것을 보장해 주는 다음과 같은 정리를 말하고 있다.

V_{FR} 를 역할 분석에 의해 생긴 텐서 곱 표상이라고 해보자. 구조 s 에서 결합된 역할을 표상하는 벡터가 선형적으로 독립적이라고 가정하자. 그러면 각 역할은 완전히 정확하게 분해될 수 있다. 곧 각 결합된 역할 r 에 대해서 s 를 표상하는 벡터 $V_{FR}(s)$ 로부터 r 에 결합된 점유자 f 를 표상하는 벡터 f 를 얻는 과정이 있다.

이 정리와 그에 대한 증명은 [4]의 186쪽을 보라. 한편 이 정리에서 벡터들이 선형적으로 독립적이지 못하다면 요소 벡터들을 되찾을 수 있는 가능성은 정도의 문제로 남는다고 Horgan과 Tiensen은 말하고 있다. 곧 선형적인 독립성이 작을수록 요소가 되는 점유자, 역할 벡터를 되찾을 수 있는 수학적 가능성은 그 정확도가 줄어든다는 것이다([18], 208쪽).

더 정확히 말하면 연결주의 모형의 활성화 벡터는 **고전적인 요소**(classical constituency)를 갖지 않는다. 포더와 맥래플린은 고전적 요소성에 대해서 이렇게 설명한다.

고전적인 견해는 정신적인 표상의 구문론이 자연 언어 문장의 구문론과 다음과 같은 점에서 같다고 주장한다: 둘 다 우리가 **고전적인 요소**라고 부를 것으로부터 만들어진 복합적인 기호들을 포함한다...우리는 이렇게 가정한다. 한 쌍의 표현 E1과 E2에 대해서 E1이 E2의 **고전적 요소**인 것은 E2가 사례화할(tokened) 때마다 E1이 사례화할 때 뿐이다([10], 333쪽).

'고전적'이라는 형용사가 붙은 요소와 그렇지 않은 요소가 갖는 차이점은 그 요소가 자신이 나타나는 표현 안에서 구체적으로 사례화하느냐 그렇지 않느냐의 차이임을 알 수 있다. 이 차이가 어떤 결과를 가져오기에 포더와 그의 동조자들은 참다운 요소가 되느냐 못되느냐의 결과까지 끌어 낼 수 있다고 생각하는가? 그것은 정신 과정의 관점에서 복합적인 기호의 고전적인 요소들은 사실상 복합적인 기호가 사례화하는데 인과적으로 영향을 끼치는데 비해 연결주의에서 말하는 텐서 곱과 중첩 벡터와 같은 요소들은 인과적인 상태 자체를 가질 수 없다는 점 때문이다. 한 마디로 말해서 구체적으로 일어난 일은 그것이 일어나 있다는 자체만으로도 무슨 일인가를 일어나게 하는 데 참여하고 있지만 단

지 가상의 것은 무슨 일이든 일어나게 할 수는 없다. 포더와 맥래플린에 따르면.

...복합 기호의 고전적인 요소들은 복합 기호가 사례화할 때 인과적인 영향을 결정하는 데 이바지한다...그러나 거칠게 말해서 육이 "3×2"로 표상될 수 있다는 사실은 그것만 가지고는 육이 "6"으로 표상되는 셋틀[컴퓨터](또는 뇌)에서 일어나는 인과적 과정에 영향을 끼칠 수 없다. 단순한 반사실적 표상들은 인과적인 영향력을 갖지 못한다. **실제로 사례화되는 표상들만이 인과적인 영향력을 가질 뿐이다** ([10], 346쪽; 강조는 글쓴이가).

정신적인 표상이 인과적인 힘을 갖기 위해서는 그 기호들이 사례화되어야 한다는 것이다.

포더와 맥래플린의 주장을 정리해 보면 다음과 같다.

1. 올바른 구성적 구조는 복합 표현에서 요소가 되는 표현이 구체적으로 사례화하여야 하고 어떤 복합 표현에서도 그 사례가 보존되어야 한다.
2. 그래야만 하는 이유는 실제로 사례화하는 요소만이 인과적인 영향력을 행사할 수 있기 때문이다.
3. 단지 가상의 것은 인과적인 힘을 가질 수 없다. 그런데 연결주의에서 말하는 요소는 가상의 것이다.

4. 그런데 실제로 사례화하는 기호적인 구조를 우리 마음도 지녀야 한다.

5. 그러므로 사고 언어의 구성적 구조는 구체적으로 사례화하는 고전적인 요소를 포함하고 있어야 하고 그런 특징을 보여 주지 못하는 연결주의 모형은 올바른 모형이 아니다.

이와 같은 지적은 텐서 곱 표상에 대해서 틀린 말은 아니다. 포더 등은 텐서 곱 표상에서 나오는 점유자와 역할 유닛의 활성화 상태가 **가상적(imaginary)**이라고 주장한 것이다([10], 344쪽). 이게 무슨 말인가 들어 보기 위해 '성자'를 표상하는 기계를 다시 생각해 보자. 낱말 '성자'를 표상하는 중첩 패턴은 기계에서 실제로 실현되는 활성화 벡터일 것이다. 그러나 글자 '성'을 표상하는 활성화 벡터는 단지 가상적이다. 또 '첫번째 위치에 있는 글자임'을 표상하는 활성화 벡터도 가상적이다. 그리고 또 '첫번째 위치에 있는 글자 "성"을 표상하는 텐서 곱 벡터도 마찬가지로 가상적이다. 가상적이지 않은 활성화 패턴은 '성자'를 표상하는 중첩 벡터일 뿐이다. 포더와 맥래플린이 이렇게 주장하는 근거는 딱 하나 뿐이다. '성자'를 표상하는 벡터 표상에서 실제로 사례화된 것은 '성자'를 표상하는 중첩 벡터 밖에 없기 때문이다. 그리고 글자 '성'이 가상의 것인 이유는 그것이 텐서 곱과 패턴 중첩을 통해서 낱말 '성자'를 표상하는 중첩 벡터를 이루게 하지만 실제로 사례화하지는 못했기 때문이다.

5. 의심 풀기

포더와 맥래플린이 텐서 곱 표상의 요소에 대해 품은 이러한 의심을 어떻게 풀어주어야 할까? 글쓴이는 두 가지 방향에서 이 의심을 풀어보려고 한다. 첫번째는 그들이 어떻게 해서 그런 의심을 가지게되었는가를 뒤쫓아가 보아, 그 의심의 근거가 희박함을 지적할 것이다. 두번째로는, 앞 절의 주장 2와 3에 주목해 본다. 과연 실제로 사례화하는 요소만이 인과적인 영향력을 행사할 수 있고, 단지 가상의 것은 인과적인 힘을 가질 수 없는가? 텐서 곱 모형의 경우를 다시 생각해 보며 그에 대한 반례를 제시해 보자.⁷⁾

첫번째로, 왜 포더와 맥래플린은 실제로 사례화되지 못한 요소는 가상적이라고 생각할까? 포더와 맥래플린은 이렇게 말하고 있다.

텐서 곱과 중첩 벡터의 요소들은 고전적 요소와 다음과 같은 점에서 차이가 난다. 곧 복합적인 고전적 기호가 개체화할 때

주7) 글쓴이는 (9)에서 이외에도 포더와 그의 동료자들이 우리 정신적 표상에 구체적으로 사례화하는 기호를 도입하게 되는 근거로서 그 기호가 있음으로 해서 우리 사고에 있다고 가정되는 인과적 측면과 의미론적 측면 사이의 연관성을 잘 드러낼 수 있다는 점을 생각하고 있다는 것을 지적했다 (117-122쪽). 지금 논의를 집중시키기 위해 이 점은 여기서 생략하자.

그것의 요소들도 개체화하는 것 말이다. [이에 견주어] 텐서 곱 벡터 또는 중첩 벡터가 개체화할 때 그것의 요소들은 (우연한 경우는 제외하고는) 개체화하지 않는다. 이 차이점이 함축하는 바는 복합적 기호의 고전적 요소들은 그것이 개체화하는 데 인과적인 영향력을 행사하지만... 텐서 곱과 중첩 벡터의 요소들은 그런 인과적인 상태를 갖지 못한다는 점이다((10), 345쪽).

여기서 고전적인 요소란 복합 표현이 사례화(개체화)할 때 동시에 사례화하는 요소를 말한다(같은 글, 333쪽). 이를테면 '성자는 귀남을 사랑한다'가 사례화할 때 마다 '성자'가 사례화하므로 뒤의 것이 앞의 것의 고전적 요소이다. 그렇다면 위 인용문에서 텐서 곱 표상에 나오는 요소와 고전적인 요소의 차이점은 사례화하였느냐 그렇지 못하느냐의 차이 뿐이다. 다시 말해서 구현되었느냐 그렇지 못하느냐의 차이 뿐이다. 이 차이점이 함축하는 바는 정신 과정에서 고전적인 요소는 자신이 나타나는 복합 표현이 사례화할 때 동시에 사례화함으로써 복합 표현이 사례화하는 데 인과적인 영향을 끼치지 만 텐서 곱의 표상들은 전혀 그런 인과적인 힘을 쓰지 못한다는 것이다. 자신이 속한 더 큰 표현에서 공간적인 위치를 한 자리 차지하고 있다는 게 그들이 말하는 인과적인 유효성의 전부인 것 같다.

이제 우리는 포더/맥래플린과 다음과 같은 문답을 '가상적으로' 주고 받는다.

물음: 왜 활성화 상태가 가상적면 안 되고 실제적일 것을 요구하는가?

대답: 실제적인 활성화 상태여야 인과적인 힘을 갖기 때문에 그렇다.

물음: 그럼 왜 인과적인 힘이 있어야 한다고 생각하는가?

대답: 우리는 "체계성 문제에 대한 고전적인 해결은 정신적인 표상들의 요소들이 인과적인 역할들을 갖는다고 전제한다. 다시 말해서 그것들은 정신 과정(mental process)에서 세력을 끼치는 범위를 제공한다"([10], 346쪽)고 말했다. 간단하게 말하자면 인과적인 힘이 있어야 되는 까닭은 그래야만 정신 과정에 세력을 끼치기 때문이다.

물음: 인과적인 힘이 있어야 된다는 것이나 정신 과정에 세력을 끼쳐야 한다는 것이나 결국은 같은 말을 되풀이한 것이 아닌가? 그럼 다시 묻는다. 왜 정신 과정에 세력을 끼쳐야 하는가?

대답: 우리는 방금 한 말에 바로 이어서 "따라서 기호의 고전적인 요소들은 복합적인 기호가 사례화할 때 인과적인 영향을 결정하는 데 이바지한다"고 말했다. 곧 정신 과정에 세력을 끼쳐야 하는 까닭은, 다시 말해서 인과적인 영향을 행사해야 하는 까닭은 사례화하기 위해서이다.

결국 우리가 알 수 있는 것은 이것뿐이다. 곧 **실제로 사례화해야만 인과적이 되고 인과적이 되면 실제로 사례화한다.** 실제로 사례화하는 표상들만이 인과적인 영향력을 가질 뿐이다.

요소 표현이 복합 표현에 인과적인 영향을 끼쳐야 함은 옳은 지적이다. 그렇지 않고서는 복합 표현과 요소 표현은 아무 관계도 없고 그 사이에는 포더와 필리신의 말을 빌리자면 "이름표를 써서 속인 허구"(17), 17쪽)일 뿐이기 때문이다. 그러나 그 인과적인 힘이라는 게 어떤 형식으로 나타나야 하는가? 포더와 그의 동조자가 이에 대해서 위 인용문에서처럼 그 인과적인 힘이라는 것은 구체적인 사례화에서만 생기는 것이라고 대답한다면 그건 고전적인 요소성이 필요한 이유에 대해 아무런 새로운 대답을 하지 못한 것이 되고 만다. 왜냐하면 고전적인 요소성이 필요한 이유를 인과적인 영향력으로 설명했고 다시 인과적인 영향력을 구체적인 사례화에만 국한시킨다면 다시 고전적인 요소성으로 돌아가는 악순환을 범할 뿐이기 때문이다.

그런 어리석음을 범하지 않으려면 포더와 그의 동조자들은 부분과 전체 사이에 구체적인 사례화 외에 다른 인과적인 관계가 있어도 된다고 허락해야만 할 것이다. 그렇지 않는다면 사례화해야만 인과적인 힘을 가질 수 있다는 주장에 대해서 새로운 근거를 대지 못하고 인과적인 힘은 사례화하는 경우에만 쓸 수 있다는 주장에서만 계속 맴도는 셈이 된다. 그렇다면 우리는 포더와 그의 동조자들이 꼭 사례화해야만 인과

적인 힘을 발휘할 수 있다는 데에 대해 설득력 있는 설명을 하지 않는 이상, 사례화하느냐 안 하느냐에 주목할 것이 아니라 요소가 되는 표현 또는 벡터가 인과적인 역할을 하느냐 하지 못 하느냐 하는 문제에 관심을 가져야 한다. 그래서 우리는 다른 인과적인 관계의 가능성을 생각할 수 있다. 포더와 맥래플린의 말마따나 정신적인 표상의 요소들이 정신 과정에 세력을 끼치는 범위를 제공하려면 정신 과정에서 그 요소들이 꼭 사례화하는 방법을 통하지 않더라도 전체에 어떤 형태로든 영향을 끼치면 될 것이다. 요소들 없이는 전체가 있을 수 없다는 것을 보여 주면 요소는 전체에 대해 영향력을 갖는다. 그런 의미에서 전체가 부분의 함수가 되는 경우도 부분이 전체에 영향을 끼치는 경우이다. 전체는 부분의 값이 변화함에 따라 그 값이 변화하고 이걸 바로 부분이 전체에 영향을 끼친다는 점이 아니고 무엇이겠는가?

이렇게 보듯이 연결주의 모형의 요소도 인과적 영향력을 행사할 수 있다. 그러나 그 요소는 고전주의에서 말하는 고전적 요소성과 같은 유형의 인과적 유효성은 아니다. 그리고 우리는 그런 고전적 요소성에 늘어붙어 인과적 유효성을 찾을 까닭이 없음을 이미 살펴보았다. (고전주의와 다른 방식으로 인과적 유효성을 지닌다는 점이 다음 절의 구현 논의에서 중요한 구실을 할 것이다.)

이제 나는 고전적 요소성과 같은 방식은 아니지만 전체 값이 부분 값의 함수라는 의미에서 연결주의의 요소가 인과적 영향력을 행사할 수

있음을 텐서 곱 표상에서 살펴보아 포더와 맥래플린의 의심을 푸는 두번째 작업을 하려고 한다.

텐서 곱 표상과 같은 연결주의에서 표상은 분석나무(parse tree)와 같은 이산적인 구조과 벡터 공간 사이의 짝짓기(mapping)라고 하는데 결국 그 표상은 어떤 고차원 공간의 점에 대응한다고 볼 수 있다([19], [20] 참조). 그런데 그 점이 공간에서 실제로 어디에 위치해 있느냐는 그 장소에서 벡터가 갖는 특정 값에 의해 결정된다. 다시 말해서 특정 내부 구조에 의해 결정된다. 표상을 일단 공간 속의 점으로 생각한다면 그 점들이 서로 어떤 관계에 있느냐, 다시 말해서 그 점들 사이의 '거리'가 얼마인가를 생각하는 것은 자연스러운 일이다. 이제 우리는 표상들 사이의 체계적인 관계가 전체는 부분의 함수라는, 곧 전체는 부분에 의해 결정된다는 사실을 눈치챌 수 있다. 그 표상의 내부 구조, 곧 다른 표상과 관련해서 어디에 위치해 있느냐는 다음 두 가지 점에 의해 결정된다. 첫째 그 요소의 본성. 쉽게 말하면 벡터 공간에서의 위치. 둘째 합성 과정의 본성. 결과적으로 표상이 텐서 곱 네트워크 안에서 작동하는 방식은 기능적인 요소 관계에 의존한다. 표상의 내부 구조, 곧 벡터 공간에서의 위치는 그 요소 관계의 직접적인 함수이므로 부분이 있기에 전체가 있게 되며 따라서 텐서 곱 표상은 부분과 전체 사이에 인과적인 관계가 있음을 알 수 있다. 그렇다면 고전적인 요소성에서처럼 구체적인 사례화만 인과적인 관계의 조건으로 인정해

서는 안되고 이는 다시 구체적인 사례화가 구성적 구조의 조건에서 꼭 필요한 조건의 항목이 될 수 없음을 말해 준다.

스몰렌스키가 든 원자 상태의 표상이 이 점을 이해하는 데 도움이 될 것이다. 원자의 상태는 추상적인 벡터 공간에 있는 벡터에 의해 표상된다고 한다. 이 벡터는 다음과 같이 결정된다. 곧 전자의 상태(‘회전’)는 한 벡터(V_s)에 의해 표상되고, 원자 속의 전자의 궤도는 다른 벡터(V_o)에 의해 표상된다. 그리고 원자 속에 위치하는 전자는 이 두 벡터들의 텐서 곱($V_s \otimes V_o$)에 의해 표상된다. “원자 전체는 각자 자신의 궤도에 위치해 있는 특정 전자를 표상하는 벡터들의 합 또는 중첩인 벡터에 의해 표상된다”((19), 227쪽, 주 13). 이 경우 원자 벡터값이 각자 자신의 궤도에 위치해 있는 각 전자 벡터값들의 합의 함수가 아니라고 말할 수 없을 것이다.

이 절의 결론은 이렇다. 한 마디로 말해서 텐서 곱 표상의 요소들은 인과적으로 무능하지 않다.

6. 고전주의의 구현이 아니라는 증거

이상의 논변이 성공적이라면 우리는 포더와 필리신의 한 쪽 뿔을 벗어날 수 있게 되리라. 그러나 예고했듯이 우리는 또 하나의 뿔에 맞닥뜨려야 하니 바로 구현의 위험이 그것이다. 연결주의자들이 포더와 필리신의 공격을 무디게 하려고 연결주의 모형이 구성적 구조를 보일

수 있다고 기껏 노력한 것이 오히려 고전주의의 물리적인 기반을 닦아주는 결과밖에 되지 못할 수도 있다. 그것 자체로도 큰 의미가 있는 성과겠지만 애초에 고전주의 이론과 경쟁이 되는. 그래서 그것을 갈음하려는 목적을 갖고 있었던 연결주의자들에게는 아직도 넘어야 할 고개가 남아 있는 것이다.

그럼 우선 ‘구현’이라는 말이 정확히 어떤 뜻으로 쓰이고 있고, 구현을 만족하는 조건은 무엇인가를 밝혀보도록 하자. 구현이라는 말이 낱말에서 쓰이는 경우는 이렇다면 ‘정의 사회 구현’과 같은 경우이다. 이 경우에는 정의 사회를 실제로 이룩하자, 현실화하자 등의 뜻일 것이다. 다소 느슨한 의미이지만 머리 속에 들어 있는 고안과 계획 단계의 것을 실제로 일이 일어나게끔 하자는 뜻이다. 포더와 그의 동조자들이 말하는 맥락도 마찬가지이다. 그들이 표상/풀이법 차원과 구현 차원을 의미있게 나누고 있다는 것은 이미 한 말이지만, 그들은 그 차원 구분을 통해서 표상/풀이법 차원에서 이루어지는 추상적인 설명이 구체적인 물리적 형태로 나타난다는 것을 생각하고 있다. 곧 구현이란 물리적인 구현을 말한다.

재래식 썸틀의 경우를 생각해 보자. 재래식 썸틀이 규칙 지배적인 기호 조작을 수행하고 있을 때(이렇다면 프롤로그 플그림(프로그램)을 돌리고 있을 때) 그 썸틀은 자료 처리와 개념적으로는 관계 없지만 인과적으로 관계있는 수계산을 수행함으로써 그렇게 한다. 그리고 이 수계산은 전기 신호를 보냄으로써 가능하다. 그

러나 이런 '하위 차원'의 기술들은 그 체계에 무슨 일이 일어났느냐에 대한 인지적 기술에 전혀 관련이 없어 보인다. 수 계산과 전기 신호를 들여다 봐서 뭘 알겠는가? 그것들은 그 풀 그림의 물리적인 한갓 구현일 뿐이다. 연결주의 모형에 대해서도 마찬가지로 해석할 뿐이다. 보통 연결주의 모형을 옹호하는 까닭으로서 그 모형이 재래식 폰 노이만 셈틀보다 더 뇌에 가깝다는 것을 든다. 연결망의 마디들이 뉴런과 비슷하고 그것들 사이의 연결은 시냅스와 비슷하다는 식으로 말이다. 그러나 우리가 고전적인 인지 개념(규칙 지배적 기호 조작 모형)을 구현하는 새로운 방식보다는 새로운 인지 개념에 관심이 있다면 뇌와 비슷하다는 그런 특징들은 문제가 되지 못한다. 그것들은 재래식 셈틀의 수 계산과 전기 신호처럼 구현 차원에서나 관심을 끄는 것이고, 정신 상태와 과정을 설명하는 인지 차원에는 끼어들지 못하는 것이다.

그러나 어느 표상/풀이법이나 어느 구현과 관계 맺는 것은 아니며 어느 구현이나 다 어느 표상/풀이법의 실현이 되는 것은 아니다. 범위를 좁혀 상위 차원과 같은 생각을 담고 있는 하위 차원의 설명이라고 해서 구현이 되는 것은 아니다. 우리는 구현의 조건을 구체적으로 알아 놓아야 하고, 그 조건에 맞는 경우 구현 관계에 있다고 말할 수 있어야 한다. 스몰렌스키는 부족함이 없게 정의를 한 것은 아니지만 이와 같은 점을 귀뜸해 줄 수 있는 다음과 같은 말을 하고 있다.

만약 한 계산 체계에 대해 어떤 차원에서 설명이 있고 그보다 하위인 다른 차원에서 설명이 있다면, 하위 차원이 상위 차원의 구현이 되는 필요 충분 조건은 상위의 기술이 그 체계의 행동에 대해 완전하고, 정확한 풀이법적 서술이 되어야 한다는 것이다 ([19], 203쪽).

상위 기술은 풀이법을 통해서 하위 기술과 완전하고, 정확한 짝맺음이 가능해야 구현이라는 말을 쓸 수가 있는 것이다.

이야기를 더 진행시켜 보자. 가령 1부터 어떤 주어진 수까지의 합을 구하는 풀이법을 생각해 보자. 이 풀이법이 목표하는 바, 1부터 어떤 주어진 수까지의 합을 실현하는 방법은 여러 가지가 있을 수 있겠지만 그 구현 중에서 바로 그 풀이법을 구현한다고 말할 수 있는 경우는 구현 차원의 각 상태들이 그 상태들간의 인과적 관계를 통해서 그 풀이법에 나타나 있는 세부 풀이법 또는 그것을 나타내는 상태들간의 형식적 또는 구문론적 관계를 정확히 반영하고 있는 때 일 것이다. 표상과 언어의 경우도 마찬가지이다. 표상들이 가지고 있는 의미론적 관계를 구현에서 인과적인 힘으로 바꾸어 그대로 나타내 줄 수 있을 때 바로 그 표상의 구현이라고 우리는 말할 수 있다. 그러기 위해서는 먼저 상위 차원을 이루는 각 상태들과 하위 차원의 상태들의 짝맺음(mapping)이 가능해야 할 것이다. 그래서 상위 인지 차원의 상태들간의 의미론적, 구문론적, 형식적 관계가 하위 구현 차원의

상태들간에서 인과적 관계를 통해 보전시켜 주어야 할 것이다.

차머스는 이를 다음과 같이 정리하고 있다.

한 물리적 체계가 주어진 계산을 구현하는 경우는 그 물리적 체계의 인과적 구조가 계산의 형식적 구조를 내비치는 때이다 ([21], 3쪽).

그는 이를 좀더 자세히 설명하고 있다.

한 물리적 체계가 주어진 계산을 구현하는 경우는 그 체계의 물리적 상태들을 상태-유형들로 때를 짓게 하는 것이 가능하고, 계산의 형식적 상태에서 그렇게 나온 물리적 상태-유형으로 짝짓기가 존재함으로써, 추상적인 상태-움김 관계가 관계를 맺어주는 형식적 상태들이 그와 상응하는 인과적인 상태-움김 관계가 관계를 맺어주는 물리적인 상태-유형과 짝을 짓게 되는 때이다 ([21], 3쪽).

단순히 상태-유형들 사이의 짝짓기만으로는 어느 계산과 어느 구현과도 짝을 지을 수 있다. 그게 바로 계산의 막강한 힘이니까. 그래서 차머스는 추상적인 상태-움김 관계가 관계를 맺어주는 형식적 상태들이 그와 상응하는 인과적인 상태-움김 관계가 관계를 맺어주는 물리적인 상태-유형과 짝을 짓게 되어야 한다는 조건을 잊지 않고 있는 듯 하다.

그렇다면 하위 구현 차원이 상위 인지 차원의 구현이 되려면 다음 두 조건을 만족해야 한다.

- ① 상위 인지 차원의 각 상태들과 하위 구현 차원의 각 상태들의 짝맺음.
- ② 그 짝맺음을 통해 상위 차원의 구문본적, 형식적, 의미론적 관계가 하위 차원에서 인과적 관계로 보존됨.

포더와 그의 동조자들은 이와 같은 구현의 조건을 너무 강한 것이라고 생각할까? 그 속내야 알 수 없지만 글쓴이 생각으로는 그렇지 않을 것 같고, 또 그럴 수도 없을 것 같다. 우리는 이미 보았듯이 포더와 그의 동조자들은 구성적인 구조에 대해 어떤 꿀바꿈에도 곳곳하게 그 모습을 바꾸지 않고 보존하는 고전적인 요소성을 요구하고 있다. 그것을 요구하는 까닭은 그래야만 인과적인 유효성을 가장 잘 발휘할 수 있기 때문이라는 것도 이미 보았다. 그런데 고전적인 요소성은 표상 차원에서 하는 얘기지만 그 얘기가 인과적인 유효성을 더 많이 필요하면 필요했지 적게 필요하지는 않을 물리적인 구현의 단계에서도 그대로 요구하리라는 것은 당연한 추측일 테고, 역시 마찬가지로 그 표상 차원의 상태-유형이 구현 차원의 상태-유형과 맞물려야 한다고 말해야 하지 않겠는가? 그들의 처지에서만 봐도 말이다.

더구나 연결주의자들과 싸우기 이전에 나온 포더의 글에서 이런 생각을 확인할 수 있다. 포

더는 “정신 상태의 인과적 역할은 정신 상태의 명제적 대상이 갖는 함축 구조(곧 의미론적, 구문론적 구조)와 전형적으로 밀접하게 유사하다”(21), 90쪽: (14), 18쪽)고 말한다. 그런데 이 인과적 관계와 의미론, 구문론적 관계를 가장 잘 보존해 주는 것이 기호를 조작하는 재래식 셈틀이라고 한다. 포더가 생각하기에 의미론적 속성과 인과적인 속성이 가장 잘 연결되는 매체는 기호이다. 그래서 “셈틀은 기호의 인과적인 속성과 의미론적 속성을 매개하는 문제의 해결책이 되는”(22), 94쪽: (14), 19쪽) 것이다.

그럼 셈틀은 어떤 특징이 있길래 그런 해결책이 될 수 있을까? 이 물음에 대해 “당신은 기호의 구문론을 통해서 인과적인 속성을 의미론적 속성과 연관지을 수 있다”(22), 93쪽: (14), 18쪽)는 구절이 답이 될 수 있다. 기호의 구문론은 이차적인 물리적 속성 가운데 하나다. 다시 말해서 기호의 구문론적 구조란 거창한 게 아니라 그 기호가 갖는 기하학적 생김새이다. 구문론이란 그런 생김새로 바꾸어 말할 수 있고 그 기호의 생김새가 인과적인 역할을 잠재적으로 결정해 주기 때문에 기호의 인과적인 역할이 구문론과 관련을 맺는 상황을 쉽게 상상할 수 있다. “열쇠의 생김새가 그 열쇠가 어느 자물쇠를 열 것인가를 결정한다”는 포더의 비유처럼(22), 93쪽: (14), 19쪽), 기호의 구문론 곧 생김새가 그 기호의 사례화의 원인과 결과를 결정한다.⁴⁾

포더는 기호를 조작하는 폰 노이만식 기계

만 알고 있을 때 구현의 관계에 대해서 이와 같이 생각했다. 그 관계를 연결주의 기계에 대해서도 똑같이 생각할 것이고 또 그래야만 할 것이다. 그렇다면 포더와 그의 동조자들이 위에서 내린 구현의 조건에 뜻을 함께 하리라고 가정하고서 연결주의 모형은 결코 고전주의 모형의 구현이 아니라는 증거를 몇 가지 대겠다.

1) 앞 절에서 포더와 맥래플린이 연결주의 모형의 요소가 갖추지 못했다고 비난했던 고전적 요소 덕분에 연결주의 모형은 오히려 구현의 덩어리에서 빠져나갈 수 있는 것 같다. 일단 앞 절의 논의가 성공이라는 전제 아래 연결주의 모형의 구성적 구조가 포더와 맥래플린이 요구하는 고전적 요소를 지니지 못해도 거기서 애초에 의도했던 인과적 유효성을 충분히 발휘하고 있다고 하자. 이미 알고 있다시피 고전적인 요소는 어떤 꿀바꿈에서도 그 꿀을 원래대로 보존하고 있는 요소이다. 그러나 연결주의 모형의 구성적 구조에서 보이는 요소는 그런 조건은 만족하지 못하고 있다. 그런데 만약 연결주의에서 설명하는 구성적 구조가 고전주의 인지 이론의 구현이 되려면 인지 차원의 고전적 요소를 구현 차원에 있는 연결주의 모형의 요소가 위에서 말한 구현의 조건에 맞게 구현해야 할 것이다. 그러나 두 요소들이 이렇듯이 중요한 점에서 차이

주8)이 부분에 대한 더욱 상세한 논의는 (9)3.4.2절을 참조하라.

를 보이고 있으니 고전적 요소와 연결주의 모형의 요소 간의 짝짓기가 애초에 가능하지가 않는 것이다. 연결주의 모형의 요소에서는 고전적 요소와 짝을 맺을 수 있는 명시적이며 보존적인 요소가 있지 않기 때문이다.

그러나 연결주의 모형의 요소가 너무 깊이, 잠재적으로 들어 있어서 그렇지 고전적 요소와 원리적으로 짝짓기가 불가능한 것은 아니라는 지적이 있을 수 있다. 그에 대해서는 다음을 보자.

2) 이제는 널리 알려져 있지만, 연결주의 모형의 표상은 분산(distributed) 표상이다. 간단한 요소들과 표상하려고 하는 존재자로 이루어진 네트워크에서 한 존재자를 표상하기 위해 한 가지 요소만을 사용하는 방식을 국부 표상이라고 하고, 반면에 각 존재자를 여러 요소에 분산된 활성화 패턴으로 표상하려는 방식을 분산 표상이라고 한다. 분산 표상에서 각 요소들은 여러 존재자를 표상하는 데 참여한다. 그래서 분산 체계와 같은 유형의 네트워크에서는 특정 표상에서 활성화되는 유닛들이 표상되는 대상의 미시 특성을 부호화한다. 분산 표상의 이런 성질 때문에 연결주의 모형의 표상은 고정적이지 못하고, 문맥 의존적이며, 가족 유사성을 잘 드러낸다는 특징을 보여주고 있다. 그리고 이는 고전주의 모형의 기호에 의한 표상이 고정적이고, 문맥 독립적이라는 특징과 잘 견주어진다. 구성적 구조에서 요소들이 보이는 문맥 의존성, 그 요소들끼리 결합하려고 할 때 생기는 상호

작용, 요소들을 정확히 유일하게 집어내려고 할 때 생기는 곤란함, 한 개념을 표상하려고 할 때, 이를테면 스몰렌스키의 '커피'를 표상하려고 할 때 그 개념이 가족 유사성에 의해 묶여지는 벡터들의 합으로 이루어져 있다는 사실, 이런 것들은 연결주의 모형에서 구성적 구조를 보인다고 할 때 그 구조가 고전주의 인지 이론의 구현이라고 하는 주장을 다시 한 번 심각하게 깊이 깊이 생각하도록 만들어 준다. 구현의 조건이 바라는 짝맺음이 가능할 것 같은가?

맥래플린은 연결주의가 결국은 뇌가 튜링의 열쇠를 어떻게 여는지 말해 줄 것이라고 추측한다((13), 165쪽). 그러나 부드럽고 연약한 연결주의적 요소로는 그 열쇠를 열 수가 없다.

3) 차머스는 구조에 민감한 처리 과정이 언제나 고전적인 요소 표상들로 분해하는 과정을 통해서만 진행되어야 한다면, 연결주의 모형은 고전주의의 기호 처리 모형의 구현에 불과하다고 말한다((2), 312-5쪽 참조). 옳은 지적이다. 연결주의 모형의 구성적 구조도 항상 고전적인 요소 표상들로 분해하는 과정을 거쳐서 진행되어야 한다면, 그것은 바로 고전적인 기호 처리 모형과 짝지을 수 있는 조건을 마련해 줄 수 있기 때문이다. 차머스는 이어서 그렇게 요소적인 부분을 뽑아내지 않고서도 일련의 꿀바꿈 과정을 실행할 수 있는 모형을 제시한다. 그 모형은 '성자는 귀남이를 폭행했다'와 같은 제힘꼴(능동형)이 '귀남이는 성자에게 폭행을 받았다'와 같은 입은꼴(수동형)로 꿀바꿈하는 작

업을 한다. 이 작업이 성공적으로 이루어짐으로써 연결주의 모형에서 구조적인 표상이 가능함을 차머스는 보이려고 한다. 그는 그러면서 동시에 고전주의 인지 이론의 구현에 기대지 않고서도 가능하다는 것을 주장하려고 한다. 세부적인 기술은 덜고 설명하면, 이 네트워크는 그 꿀바꿈 과정에서 '성자', '귀남', '폭행하기'와 같은 명시적인 요소를 거치지 않는다. 물론 그것이 가능한 것은 39-19-39 개의 마디들로 이루어진 네트워크가 2)에서 말한 것처럼 분산된 형태로 표상을 가능하게 했기 때문이다. 한 마디로 말해서 차머스가 말하는 네트워크는 문장의 제힘꼴을 입은꼴로 바꾸는 꿀바꿈 과정에서 명시적인 고전적인 요소로 분해되는 과정을 꼭 통하지 않고도 그의 말대로 '직접적이고 전체론적인' 방법을 통해 가능했기 때문에 구현의 짐을 벗어버릴 수 있다는 것이다.

차머스의 주장에서 오해하지 말아야 할 것은 꿀바꿈 과정에서 요소로 분해하는 과정이 전혀 필요없다는 것은 아니라는 점이다. 그 요소가 명시적으로 사례화하는 고전적인 요소일 필요가 없다는 것이다. 이 네트워크의 요소는 그런 요소보다 더 미세하고 네트워크 전체에 분산되어있는 요소인 것이다.

차머스의 이런 주장을 스몰렌스키의 텐서 곱 표상에도 그대로 부려 쓸 수 있겠다. 앞 절의 텐서 곱 표상은 한 표상에 요소 구조가 있음을 연결주의 모형이 보여줄 수 있다는 것이었으므로 꿀바꿈을 보이는 것은 아니었다. 꿀바꿈을 살펴보기 위해 그 모형에서 체계성을 어떻게 보

일 수 있는가를 알아 보자. 텐서 곱 표상은 표상의 체계성을 확실히 보여 줄 수 있다. 한 문장의 텐서 곱 표상은 한 구문론적 역할을 표상하는 벡터와 그 역할을 맡는 점유자를 표상하는 벡터를 곱하고, 그 문장의 각 요소들에 대해 그렇게 해서 생긴 텐서 곱들을 중첩해서 생긴다고 했다. 같은 유형의 표상들을 '성자'와 '귀남'이라고 하고, 그 유형의 표상이 맡을 수 있는 역할을 '사랑하기'라고 해 보자. 만약 $f_{성자} \otimes r_{사랑하기}$ 가 어떤 네트워크에서 역할 '사랑하기'에 있는 '귀남'의 표상이라면, $f_{성자} \otimes r_{사랑하기}$ 는 그 네트워크에서 역할 '사랑하기'에 있는 '성자'의 표상일 것이다. 일반적으로 여러 역할들을 맡을 수 있는 유형들의 각각의 역할과 점유자의 벡터 표상들이 주어진다면, 한 점유자가 한 역할을 점유할 수 있다면 같은 유형의 다른 어떤 점유자도 그 역할을 맡을 수 있다.⁹⁾ 따라서 '성자는 귀남이를 사랑한다'를 표상하는 것이 가능한 네트워크라면 '귀남이는 성자를 사랑한다'를 표상하는 것도 가능하다. 텐서 곱 표상은 표상의 체계성을 이렇게 보여 주고 있다.

이렇게 볼 때 텐서 곱 표상에서 요소적인 표상이 체계적으로 관련을 맺고 있다는 것이 전부 다이다. 그 요소가 고전적인 요소일 필요는 없다. 그러므로 차머스의 제힘꼴-입은꼴 꿀바꿈 모형에서처럼 텐서 곱 표상에서도 고전적인 명시적 요소를 거치지 않고서도 체계성이 가능

주9)[18]. 213쪽 참조.

함을 보일 수 있는 것이다.

7. 맺는말

글쓴이는 5절에서 고전적인 요소성이 지니는 개념적인 난맥상을 지적했다. 6절에서는 고전적인 요소성을 통해서 연결주의 모형이 포더와 펠리신의 구현 조건을 벗어날 수 있는 길을 더듬어 찾아 보았다. 구현 조건이 스몰렌스키나 차머스가 말하는 식의 조건이라면 고전적 요소성은 오히려 고전주의 인지 이론의 연결주의적 구현을 방해하는 것이 될 것이다. 고전주의 모형의 각 상태의 구문론적 관계가 연결주의 모형의 각 상태의 인과론적 관계를 반영하여 행복한 짜릿기를 이루기에는 고전주의 모형에 고전적 요소성이라는 강한 가시가 뚫혀 있는 것이다.

그렇다면 포더와 그의 동조자들 처지에서는 연결주의 모형이 꼭 자신들의 인지 이론의 구현이 되지 않더라도 어쨌든 인지 이론과 상관 없는 물리적인 실현의 이론이라고 말할 수도 있겠다. 어떤 상위의 인지 이론을 구현하는지는 모르지만 말이다. 그러나 이런 반응은 그들의 애초의 주장과 다른 길로 가는 것이다. 그들은 연결주의의 매력을 대하면서 그것을 인정하며 그것과 화해하여 자신들의 이론에 주는 위협을 줄이기 위해 연결주의 이론이 고전주의 이론의 구현이라고 말한 것이었다. 그런데 그런 식의 구현 공격이 제대로 먹혀 들어가지 않으니 아예 우리 이론과는 전혀 상관 없는 물리적인 구현

의 이론이라고 말할 수도 있겠다. 그럴 수도 있겠다. 그러나 그렇다면 고전주의자들은 사변적인 논의 말고 자신들의 이론을 뒷받침해 줄 물리적인 구현 차원의 경험적인 이론을 따로 찾아 나서야 한다. 고전주의의 구현 이론이 연결주의자들이 즐겨 인용하는 신경 생리학의 연구 결과와 많은 부분에서 어긋나는 데야 어쩔 수 없을 것이다. 더 나아가 '구현 연결주의자'라고, 물리적인 실현의 연결주의에만 관심있는 연결주의자들이 아니고서야 자신들의 이론이 표상과 인지에 대해서 말하고 있다고 생각하지 않는 연결주의자들은 드물 것이다. 그런 생각이 연결주의자들 자신들의 생각만으로 그친다면야 모르겠지만, 우리는 이미 연결주의 모형이 표상 차원에서 말하는 바가 고전주의 인지 이론에 위협을 주는 경쟁적인 대안을 많이 알고 있지 않은가? 그 때 가서 또 다시 그것은 우리 이론의 구현 차원의 이론일 뿐이라고 말할 것인가?¹⁰⁾

주10)이 글의 애벌린고는 1994년 9월 9일 소흥렬 교수님 연구실에서 모인 HCI 기초이론부 모임과 1995년 2월 17일 잠실 롯데호텔에서 열린 HCI 학술대회에서 읽었다. 그 때 좋

참고문헌

- [1]Chalmers, D.J.(1990), "Syntactic Transformations on Distributed - Representations", *Connection Science* 2: 53-62.
- [2]Chalmers, D.J.(1993), "Connectionism and Compostionality: Why Fodor and Pylyshyn Were Wrong", *Philosophical Psychology* 6: 305-319.
- [3]Smolensky, P.(1987), "The Constituent Structure of Connectionist Mental States: A Reply to Fodor and Pylyshyn" in T. Horgan and J. Tien- sen(eds.)(1991), *Connectionism and the Philosophy of Mind*, Dordrecht: Kluwer Academy Publishers: 281-308.
- [4]Smolensky, P.(1990), "Tensor Product Variable Binding and the Representation of Symbolic Structures in Connectionist Systems", *Artificial Intelligence* 46: 159-216.
- [5]Clark, A.(1989), *Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing*. Cambridge, MA: MIT Press/Bradford Books.
- [6]van Gelder, T.(1990), "Compositionality: A Connectionist Variation on a Classical Theme", *Cognitive Science* 14, 353-384.
- [7]Elman, J.L.(1990), "Structured Representations and Connectionist Models", in G. Altmann(ed.), *Cognitive Models of Speech Processing: Computational and Psycholinguistic Perspectives*, Cambridge, MA: MIT Press/Bradford Books.
- [8]Pollack, J.B.(1990), "Recursive Distributed Representation", *Artificial Intelligence* 46: 77-105.
- [9]최 훈(1994), 《연결주의와 구성적 구조: 포더와 그의 동조자들에게 보내는 응답》, 서울대학교 철학과 석사학위 논문. (이 논문의 일부가 "연결주의와 구성적 구조"라는 제목으로 《철학논구》 제22집, 서울, 서울대학교 철학과: 229-276쪽에 실림).
- [10]Fodor, J.A and B.P. Mc-

은 도움 말씀을 주신 소홍렬 교수님, 공용현 박사님, 이영의 선생님과 그의 분들에게 감사드린다. 그리고 초고를 꼼꼼히 읽고 많은 부분을 바로 잡아주신 김영정 교수님께도 감사드린다.

- Laughlin(1991), "Connectionism and the Problem of Systemicity: Why Smolensky's Solution Doesn't Work", in T. Horgan and J. Tiensen(eds.)(1991), *Connectionism and the Philosophy of Mind*, Dordrecht: Kluwer Academy Publishers: 331-354.
- [11]Fodor, J.A. and Z.W. Pylyshyn(1988), "Connectionism and Cognitive Architecture: A Critical Analysis", *Cognition* 28: 3-71.
- [12]Butler, K.(1993), "Connectionism, Classical Cognitivism and the Relation between Cognitive and Implementational Levels of Analysis", *Philosophical Psychology* 6: 321-333.
- [13]McLaughlin, B.(1993), "The Connectionism/Classicism Battle to Win Souls", *Philosophical Studies* 71: 163-190.
- [14]Fodor, J.(1987), *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*, Cambridge, MA: MIT Press/Bradford Books.
- [15]최 훈(1993), "라이벌이냐 구현이냐?", HCI 연구실.기초이론부(편역), 《HCI 기초이론》, 서울: 29-44쪽.
- [16]Rumelhart, D.E. and J.L. McClelland(1986), "PDP Models and General Issues in Cognitive Science", in D.E. Rumelhart and J.L. McClelland and the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructures of Cognition, vol. 1: Foundations*, Cambridge, MA: MIT Press/Bradford Books: 109-146.
- [17]McClelland, J.L. and Rumelhart, D.E.(1981), "An Interactive Activation Model of Context Effects in Letter Perception: Part 1", *Journal of Experimental Psychology: General*, 114: 159-188.
- [18]Horgan, T. and J. Tienson(1992), "Structured Representations in Connectionist Systems", in S. Davis(ed.), *Connectionism: Theory and Practice*, NY: Oxford University Press: 195-228.
- [19]Smolensky, P.(1991), "Connectionism, Constituency, and the Language of Thought", in B. Loewer and G. Rey(eds.)(1991), *Meaning in Mind: Fodor and his Critics*, Oxford: Basil Blackwell: 201-228.
- [20]Churchland, P.M.(1989), *A Neurocomputational Perspective: The Nature of Mind and the Structure of*

Science, Cambridge, MA: MIT Press/Bradford Books.

[21]Chalmers, D.J.(1994), "A Computational Foundations for the Study of Cognition", Manuscript.

[22]Fodor, J.(1985), "Fodor's Guide to Mental Representation: The Intelligent Auntie's Vade-Mecum", *Mind* 94: 76-100.